Journal of Statistical Planning and Inference 206 (2020) 100–110

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

# Orthogonal uniform composite designs

Xue-Ru Zhang, Min-Qian Liu, Yong-Dao Zhou *

*School of Statistics and Data Science, LPMC & KLMDASR, Nankai University, Tianjin 300071, China*

ARTICLE INFO

*Article history:*
Received 28 February 2019
Received in revised form 21 August 2019
Accepted 21 August 2019
Available online 25 September 2019

*MSC:*
62K20
62K15

*Keywords:*
Central composite design
Maximin distance criterion
Orthogonal-array composite design
Robustness
Uniform design

ABSTRACT

Composite designs are frequently utilized for fitting response surfaces in practice. This paper proposes a new type of composite designs, orthogonal uniform composite designs (OUCDs), which combine orthogonal arrays and uniform designs. Such designs not only inherit the advantages of orthogonal-array composite designs such as high estimation efficiencies and ability for multiple analysis for cross validation, but also have more flexible run sizes than central composite designs and orthogonal-array composite designs. Moreover, OUCDs are more robust than other types of composite designs under certain conditions. Some construction methods for OUCDs under the maximin distance criterion are provided and their properties are also studied. It is shown that many constructed OUCDs are maximin distance designs.



## 1. Introduction

The response surface methodology, proposed by Box and Wilson (1951), is widely applied to explore the unknown relationship between explanatory variables and interesting responses. Second-order models can be used to fit such a nonlinear relationship. A design is called a second-order design if it can be used to fit the second-order model. Several types of second-order designs have been proposed in the literature, such as central composite designs (CCDs) proposed by Box and Wilson (1951), small composite designs (Draper and Lin, 1990), subset designs (Gilmour, 2006), augmented pairs designs (Morris, 2000), definitive screening composite designs (DSCDs) proposed by Zhou and Xu (2017) and orthogonal-array composite designs (OACDs) introduced by Xu et al. (2014). Among them, OACDs which combine two-level and three-level orthogonal arrays (OAs) have appealing properties, for example, they have higher $D$-efficiencies than many other types of designs under second-order models and can perform separate analysis for the two-level OAs and three-level OAs. When second-order models are insufficient to describe the relationship between variables of importance and responses, Zhang et al. (2018b) studied OACDs which combine two-level and four-level OAs and can be used for fitting third-order models. However, the three-level or four-level OAs in the OACDs often have large number of runs.

This paper proposes a new type of composite designs, called orthogonal uniform composite designs (OUCDs), which combine two-level OAs and uniform designs (UDs), to provide more flexible run sizes than OACDs and still keep the good properties of OACDs. Roughly speaking, OUCDs replace the three-level or four-level OAs in OACDs by UDs. The main idea of UDs is to scatter design points uniformly in the experimental region (Fang et al., 2018). Discrepancy is often used to measure the uniformity of designs, such as the warp-around $L_2$-discrepancy (Hickernell, 1998) and mixture discrepancy (MD) proposed by Zhou et al. (2013). The MD can overcome the shortcomings of other discrepancies and is employed in

* Corresponding author.
*E-mail address:* ydzhou@nankai.edu.cn (Y.-D. Zhou).

https://doi.org/10.1016/j.jspi.2019.08.007
0378-3758/

this paper. It will be shown that OUCDs are robust under certain conditions. We will discuss $D$-efficiencies, $T$-efficiencies and their lower bounds for OUCDs under second-order models. Moreover, space-filling properties of OUCDs under the maximin distance criterion (Johnson et al., 1990) will be investigated, and the corresponding construction methods for maximin OUCDs will be provided. Furthermore, OUCDs will be compared with other types of composite designs such as OACDs, CCDs and DSCDs. The run sizes in OUCDs are more flexible than other types of composite designs. It will be shown that OUCDs always have larger $L_1$-distances and larger estimation efficiencies than CCDs. The two different parts of an OUCD can be used for cross validation.

The remainder of this paper is organized as follows. In Section 2, the definition of OUCDs and some examples are provided. Section 3 shows the appealing properties of OUCDs in terms of robustness, maximin distance criterion and estimation efficiency. Section 4 compares OUCDs with other composite designs such as CCDs, OACDs and DSCDs, as well as UDs. Both empirical and theoretical results are provided in this section. The concluding remarks are provided in Section 5, and the proofs of theorems and propositions are shown in Appendix A. All the detailed designs are given in the Supplementary Material.

## 2. Orthogonal uniform composite designs

Let the experimental region $\mathcal{X}$ be $[-1, 1]^k$. Let OA$(N, s^k, t)$ be an orthogonal array (OA) with $N$ runs, $k$ factors, $s$ levels and strength $t$. In general, $t$ is omitted when an OA has the strength $t = 2$. An OUCD has three parts: (i) $n_1$ cube points $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ik})$ with all $x_{ij} = 1$ or $-1$, $i = 1, \ldots, n_1, j = 1, \ldots, k$, denoted by $d_1$; (ii) $n_2$ additional points with all $\boldsymbol{x}_i \in \mathcal{X}$ which are uniformly scattered in $\mathcal{X}$, $i = 1, \ldots, n_2$, denoted by $d_2$; (iii) $n_0$ center points with all $\boldsymbol{x}_i = (0, \ldots, 0)$, $i = 1, \ldots, n_0$, denoted by $d_0$. In comparison, a CCD combines $d_1$, some center points and the $n_2 = 2k$ axial points with one of $x_i = \alpha$ or $-\alpha$ and the other $x_i = 0$; an OACD combines $d_1$, some center points and a high-level OA in $\mathcal{X}$. Therefore, the most difference among OUCD, CCD and OACD is the part of $d_2$.

A two-level regular or nonregular OA, such as a full factorial or an OA with strength 2 or higher, is chosen as $d_1$ in an OUCD. To decrease the run size of $d_1$, Draper and Lin (1990) recommended to use the Plackett–Burman designs. One choice of $d_1$ is to choose a best design under the generalized minimum aberration which was proposed by Xu and Wu (2001). It was shown that generalized minimum aberration designs minimize the overall contamination of nonnegligible interactions on the estimation of main effects and tend to be model-robust under traditional model dependent efficiency criteria, see Cheng et al. (2002) and Xu et al. (2004). Another choice of $d_1$ is to find a best OA under the maximin distance criterion proposed by Johnson et al. (1990). Since the run sizes of UDs can be chosen flexibly, we use a UD as the second part $d_2$ in the OUCD such that the total number of runs in the OUCD can be less than that of the OACD.

Let U$(N, s_1, \ldots, s_k)$ be a $k$-factor U-type design in which each of the $s_j$ levels, $\{1, \ldots, s_j\}$, appears $N/s_j$ times in the $j$th column, $j = 1, \ldots, k$. Denoted by U$(N, s^k)$ when $s_1 = \cdots = s_k = s$. Then, a uniform U-type design under MD can be used as $d_2$ and its number of levels is more flexible. When $s = N$, as a special type of U-type designs, a good lattice point (GLP) set is a suitable choice for $d_2$ due to its simple structure and good uniformity. A GLP set $D = (x_{ij})$ with $N$ runs and $k$ columns is determined by the generator vector $\boldsymbol{h} = (h_1, \ldots, h_k)$ with each $h_i \in \{1, \ldots, N\}$ being coprime to $N$, where $x_{ij} = ih_j \pmod N$, $i = 1, \ldots, N, j = 1, \ldots, k$, and the multiplication operation modulo $N$ is modified so that the result falls into $[1, N]$. Moreover, the leave-one-out GLP method is often used to improve the uniformity. A leave-one-out GLP set is obtained by deleting the last row of a GLP set. All points in a U-type design $D$ are transformed into $\mathcal{X} = [-1, 1]^k$ by mapping: $f : \ell \to -1 + 2(\ell - 1)/(s_j - 1)$ for the level $\ell$ in the $j$th column of $D$ with $\ell = 1, \ldots, s_j$ and $j = 1, \ldots, k$. Denote the transformed design by $D^*$. Moreover, the center points in an OUCD are useful to estimate the pure error. Combining $d_1$ and $d_2$ to form an OUCD, we have $k!$ distinct combinations in terms of column alignments.

**Example 1.** Consider two OUCDs. The first one is the 3-factor OUCD illustrated in Table 1 which combines an OA$(8, 2^3, 3)$, the 5-run GLP set with generator vector $\boldsymbol{h} = (1, 2, 4)$, and 3 center points. The second one is the 9-factor OUCD illustrated in Table 2 which combines the $2^{9-3}$ design with $I = 127 = 348 = 13569$, the 10-run leave-one-out GLP set with generator vector $\boldsymbol{h} = (2, \ldots, 10)$, and 4 center points. The $d_1$ in the second design is not a minimum aberration OA but a maximin OA among all regular $2^{9-3}$ designs.

If we decrease the strength of the two-level parts, the run sizes of the OUCDs in Example 1 can be reduced. The run size of $d_2$ can also be changed. Then, OUCDs have more flexibility in terms of the number of runs since UD is more flexible. More OUCDs are listed in Table 3. Furthermore, an OUCD provides more information than a CCD for a second-order model, since $d_2$ in an OUCD can be used for estimating the bilinear terms, while the axial points in a CCD cannot.

## 3. Properties of OUCDs

### 3.1. Robustness

Fixed the two-level part $d_1$ and the number of center points $n_0$, we now select the optimal $d_2$ in the composite design in the sense of robustness. Consider the following generalized linear model,

$$y = \sum_{j=1}^{p} g_j(\boldsymbol{x})\beta_j + \varepsilon, \tag{1}$$

**Table 1**
A 3-factor OUCD.

| Part | Factor | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| $d_1$ | 1 | 1 | 1 |
| | 1 | 1 | −1 |
| | 1 | −1 | 1 |
| | 1 | −1 | −1 |
| | −1 | 1 | 1 |
| | −1 | 1 | −1 |
| | −1 | −1 | 1 |
| | −1 | −1 | −1 |
| $d_2$ | −1 | $-\frac{1}{2}$ | $\frac{1}{2}$ |
| | $-\frac{1}{2}$ | $\frac{1}{2}$ | 0 |
| | 0 | −1 | $-\frac{1}{2}$ |
| | $\frac{1}{2}$ | 0 | −1 |
| | 1 | 1 | 1 |
| $d_0$ | 0 | 0 | 0 |
| | 0 | 0 | 0 |
| | 0 | 0 | 0 |

**Table 2**
A 9-factor OUCD.

| Part | Factor | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $d_1$ | $2^{9-3}$ with $I = 127 = 348 = 13569$ | | | | | | | | |
| $d_2$ | $-\frac{7}{9}$ | $-\frac{5}{9}$ | $-\frac{3}{9}$ | $-\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{3}{9}$ | $\frac{5}{9}$ | $\frac{7}{9}$ | 1 |
| | $-\frac{3}{9}$ | $\frac{1}{9}$ | $\frac{5}{9}$ | 1 | −1 | $-\frac{5}{9}$ | $-\frac{1}{9}$ | $\frac{3}{9}$ | $\frac{7}{9}$ |
| | $\frac{1}{9}$ | $\frac{7}{9}$ | −1 | $-\frac{3}{9}$ | $\frac{3}{9}$ | 1 | $-\frac{7}{9}$ | $-\frac{1}{9}$ | $\frac{5}{9}$ |
| | $\frac{5}{9}$ | −1 | $-\frac{1}{9}$ | $\frac{7}{9}$ | $-\frac{7}{9}$ | $\frac{1}{9}$ | 1 | $-\frac{5}{9}$ | $\frac{3}{9}$ |
| | 1 | $-\frac{3}{9}$ | $\frac{7}{9}$ | $-\frac{5}{9}$ | $\frac{5}{9}$ | $-\frac{7}{9}$ | $\frac{3}{9}$ | −1 | $\frac{1}{9}$ |
| | −1 | $\frac{3}{9}$ | $-\frac{7}{9}$ | $\frac{5}{9}$ | $-\frac{5}{9}$ | $\frac{7}{9}$ | $-\frac{3}{9}$ | 1 | $-\frac{1}{9}$ |
| | $-\frac{5}{9}$ | 1 | $\frac{1}{9}$ | $-\frac{7}{9}$ | $\frac{7}{9}$ | $-\frac{1}{9}$ | −1 | $\frac{5}{9}$ | $-\frac{3}{9}$ |
| | $-\frac{1}{9}$ | $-\frac{7}{9}$ | 1 | $\frac{3}{9}$ | $-\frac{3}{9}$ | −1 | $\frac{7}{9}$ | $\frac{1}{9}$ | $-\frac{5}{9}$ |
| | $\frac{3}{9}$ | $-\frac{1}{9}$ | $-\frac{5}{9}$ | −1 | 1 | $\frac{5}{9}$ | $\frac{1}{9}$ | $-\frac{3}{9}$ | $-\frac{7}{9}$ |
| | $\frac{7}{9}$ | $\frac{5}{9}$ | $\frac{3}{9}$ | $\frac{1}{9}$ | $-\frac{1}{9}$ | $-\frac{3}{9}$ | $-\frac{5}{9}$ | $-\frac{7}{9}$ | −1 |
| $d_0$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

where $\boldsymbol{x} \in \mathcal{X} = [-1, 1]^k$, $g_1, \ldots, g_p$ are specified and linearly independent functions, $\beta_1, \ldots, \beta_p$ are regression parameters and $\varepsilon$ has mean 0 and variance $\sigma^2$. If the model is true, we can estimate the regression parameters $\beta_1, \ldots, \beta_p$ in (1) by some optimal designs. However, there may exist some misspecifications from the generalized linear model (1) in many cases, and one may need a robust composite design.

Suppose the true model is

$$y = \sum_{j=1}^{p} g_j(\boldsymbol{x})\beta_j + h(\boldsymbol{x}) + \varepsilon, \tag{2}$$

where $h$ is an unknown function from a class $\mathcal{H}$. Assume $h$ is orthogonal to each $g_j$, i.e., $\int_{\mathcal{X}} g_j(\boldsymbol{x})h(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 0$ for $j = 1, \ldots, p$, $\mathcal{H}$ is a reproducing kernel Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and a reproducing kernel $\mathcal{K}(\boldsymbol{x}, \boldsymbol{\omega})$ which is a real-valued function defined on $\mathcal{X}^2 = \mathcal{X} \times \mathcal{X}$ satisfying two properties: (i) $\mathcal{K}(\boldsymbol{x}, \boldsymbol{\omega}) = \mathcal{K}(\boldsymbol{\omega}, \boldsymbol{x})$, for any $\boldsymbol{x}, \boldsymbol{\omega} \in \mathcal{X}$, and (ii) $\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j \mathcal{K}(\boldsymbol{x}_i, \boldsymbol{x}_j) \geq 0$, for any $c_i \in \mathbb{R}$, $\boldsymbol{x}_i \in \mathcal{X}$, $n = 1, 2, \ldots$. The inner product $\langle \cdot, \cdot \rangle$ of the reproducing kernel Hilbert space $\mathcal{H}$ depends on a reproducing kernel $\mathcal{K}(\cdot, \cdot)$ and satisfies two conditions: (i) $\mathcal{K}(\cdot, \boldsymbol{x}) \in \mathcal{H}$, for all $\boldsymbol{x} \in \mathcal{X}$, and (ii) $h(\boldsymbol{\omega}) = \langle h, \mathcal{K}(\cdot, \boldsymbol{\omega}) \rangle$, for every $h \in \mathcal{H}$ and all $\boldsymbol{\omega} \in \mathcal{X}$. Let $d$ be an $n$-run design in $\mathcal{X}$, $\boldsymbol{x}_i \in d$, $\boldsymbol{y}$ be the column vector of response, and $\boldsymbol{g}(\boldsymbol{x}_i) = (g_1(\boldsymbol{x}_i), \ldots, g_p(\boldsymbol{x}_i))^T$, $i = 1, \ldots, n$. Under the generalized linear model (1), the least square estimator of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ is $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$, where $\boldsymbol{X} = (\boldsymbol{g}(\boldsymbol{x}_1), \ldots, \boldsymbol{g}(\boldsymbol{x}_n))^T$ is the design matrix. If there exists the misspecification $h$ and the true model is (2), then $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{h} + \boldsymbol{\varepsilon}$ where $\boldsymbol{h} = (h(\boldsymbol{x}_1), \ldots, h(\boldsymbol{x}_n))^T$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^T$. Let the loss function $L(d, h)$ be the integrated mean squared error in $\mathcal{X}$, then under model (2),

$$L(d,h) = \int_{\mathcal{X}} E[\hat{\boldsymbol{\beta}}^T \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\beta}^T \boldsymbol{g}(\boldsymbol{x})]^2 \mathrm{d}\boldsymbol{x} = \sigma^2 \mathrm{tr}\big((\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{G}\big) + \boldsymbol{h}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{G}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{h}, \tag{3}$$

where $\boldsymbol{G} = \int_{\mathcal{X}} \boldsymbol{g}(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{x})^T \mathrm{d}\boldsymbol{x}$. When $\sigma^2 = 0$, $L(d, h)$ only depends on the misspecification part $h$. If there is no misspecification from the generalized linear model (1), $L(d, h)$ only depends on the variance part $\sigma^2 \mathrm{tr}\big((\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{G}\big)$.

Yue and Hickernell (1999) gave an upper bound of the loss function $L(d, h)$ for a design $d$. We extend it to a composite design $d$ when $\beta$ is estimable, $h \in \mathcal{H}$. Let $d = (d_f^T, d_u^T)^T$, where the $n_f$-run part $d_f$ is fixed. Let the corresponding design matrix of the generalized linear model (1) be $\boldsymbol{X} = (\boldsymbol{X}_f^T, \boldsymbol{X}_u^T)^T$. Let $\boldsymbol{k}_f$ and $\boldsymbol{k}_u$ be the vectors of $n_f$ and $n_u$ kernel functions based on $d_f$ and $d_u$, respectively, i.e., $\boldsymbol{k}_\ell = (\mathcal{K}(\cdot, \boldsymbol{x}_1^\ell), \ldots, \mathcal{K}(\cdot, \boldsymbol{x}_{n_\ell}^\ell))^T$ where $\boldsymbol{x}_1^\ell, \ldots, \boldsymbol{x}_{n_\ell}^\ell \in d_\ell$, and $\boldsymbol{K}_{\ell s} = \boldsymbol{k}_\ell \boldsymbol{k}_s^T$, $\ell, s = f, u$.

**Theorem 1.** *For the design matrix $\boldsymbol{X} = (\boldsymbol{X}_f^T, \boldsymbol{X}_u^T)^T$, suppose that $\boldsymbol{\beta}$ is estimable, then*

$$L(d,h) \le \gamma(d,h) = \sigma^2 \lambda_{\min}^{-1}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right) \mathrm{tr}\left(\boldsymbol{G}\right) + \lambda_{\max}\left(\boldsymbol{G}\right) \lambda_{\min}^{-2}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right) \left(\lambda_{\max}\left(\boldsymbol{X}_f^T\boldsymbol{K}_{ff}\boldsymbol{X}_f\right) + \lambda_{\max}\left(\boldsymbol{Q}\right) + \lambda_{\max}\left(\boldsymbol{X}_u^T\boldsymbol{K}_{uu}\boldsymbol{X}_u\right)\right) \|h\|^2, \tag{4}$$

*where $\boldsymbol{Q} = \boldsymbol{X}_f^T\boldsymbol{K}_{fu}\boldsymbol{X}_u + \boldsymbol{X}_u^T\boldsymbol{K}_{uf}\boldsymbol{X}_f$, $\lambda_{\max}(\boldsymbol{Y})$ and $\lambda_{\min}(\boldsymbol{Y})$ are the maximum and minimum eigenvalues of any square matrix $\boldsymbol{Y}$, respectively. The equality holds when $d_u$ is empty, $\boldsymbol{X}_f^T\boldsymbol{X}_f = a\boldsymbol{I}_p$, $\boldsymbol{G} = b\boldsymbol{I}_p$, $h = c\boldsymbol{v}_{\max}^T(\boldsymbol{T})\boldsymbol{G}^{1/2}(\boldsymbol{X}_f^T\boldsymbol{X}_f)^{-1}\boldsymbol{X}_f^T\boldsymbol{k}_f$, where $a, b, c \in \mathbb{R}$, $a, b > 0$, and $\boldsymbol{v}_{\max}(\boldsymbol{T})$ is the (scaled) eigenvector corresponding to the maximum eigenvalue of $\boldsymbol{T} = \boldsymbol{G}^{1/2}(\boldsymbol{X}_f^T\boldsymbol{X}_f)^{-1}\boldsymbol{X}_f^T\boldsymbol{K}_{ff}\boldsymbol{X}_f(\boldsymbol{X}_f^T\boldsymbol{X}_f)^{-1}\boldsymbol{G}^{1/2}$.*

It is difficult to search a robust $n$-run exact design because of its discrete structure. Similar to the continuous design theory in Kiefer (1959), we consider the continuous design measure. For a design with $n$ runs and the probability distribution $\xi$, define $\boldsymbol{X}^T\boldsymbol{X} = n\int_{\mathcal{X}} \boldsymbol{g}(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{x})^T \mathrm{d}\xi(\boldsymbol{x})$, and $\boldsymbol{H}(\xi) = \boldsymbol{X}^T\boldsymbol{K}\boldsymbol{X} = n^2\int_{\mathcal{X}\times\mathcal{X}} \boldsymbol{g}(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{\omega})^T\mathcal{K}(\boldsymbol{x}, \boldsymbol{\omega})\mathrm{d}\xi(\boldsymbol{x})\mathrm{d}\xi(\boldsymbol{\omega})$. Assume that $d_u$ has the probability distribution $\xi_u$, and denote the corresponding composite design by $d_{\xi_u}$. Denote $\boldsymbol{H}_{uu}(\xi_u) = n_u^2\int_{\mathcal{X}\times\mathcal{X}} \boldsymbol{g}(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{\omega})^T\mathcal{K}(\boldsymbol{x}, \boldsymbol{\omega})\,\mathrm{d}\xi_u(\boldsymbol{x})\mathrm{d}\xi_u(\boldsymbol{\omega})$, $\boldsymbol{H}_{fu}(\xi_u) = \boldsymbol{X}_f^T\boldsymbol{K}_{fu}\boldsymbol{X}_u = n_u\boldsymbol{X}_f^T\int_{\mathcal{X}}\langle\boldsymbol{k}_f, \mathcal{K}(\cdot, \boldsymbol{\omega})\rangle\boldsymbol{g}(\boldsymbol{\omega})^T\mathrm{d}\xi_u(\boldsymbol{\omega})$ and $\boldsymbol{H}_{uf}(\xi_u) = \boldsymbol{H}_{fu}^T(\xi_u)$. Next, we want to find the best $d_u$ with the probability distribution $\xi_u$ such that the composite design $d_{\xi_u}$ is robust. The following theorem shows that the design $d_u$ with the uniform distribution $\xi_u$ is most robust when $d_f$ is fixed.

**Theorem 2.** *For any composite design $d_{\xi_u}$, the upper bound (4) becomes*

$$\gamma(d_{\xi_u}, h) = \sigma^2\lambda_{\min}^{-1}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)\mathrm{tr}(\boldsymbol{G}) + \lambda_{\max}\left(\boldsymbol{G}\right)\lambda_{\min}^{-2}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)\left(\lambda_{\max}\left(\boldsymbol{X}_f^T\boldsymbol{K}_{ff}\boldsymbol{X}_f\right) + \lambda_{\max}\left(\boldsymbol{H}^S(\xi_u)\right) + \lambda_{\max}\left(\boldsymbol{H}_{uu}(\xi_u)\right)\right)\|h\|^2, \tag{5}$$

*where $\boldsymbol{H}^S(\xi_u) = \boldsymbol{H}_{fu}(\xi_u) + \boldsymbol{H}_{uf}(\xi_u)$. If $\boldsymbol{H}^S(\xi_u)$ is not a negative definite matrix, $\gamma(d_{\xi_u}, h)$ can reach the minimum value $\gamma_{\min}(d_{\xi_u}, h)$ when $\xi_u$ is the uniform measure in $\mathcal{X}$, namely,*

$$\gamma_{\min}(d_{\xi_u}, h) = \sigma^2\lambda_{\min}^{-1}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)\mathrm{tr}(\boldsymbol{G}) + \lambda_{\max}\left(\boldsymbol{G}\right)\lambda_{\min}^{-2}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)\lambda_{\max}\left(\boldsymbol{X}_f^T\boldsymbol{K}_{ff}\boldsymbol{X}_f\right)\|h\|^2. \tag{6}$$

From Theorem 2, $\gamma_{\min}(d_{\xi_u}, h)$ can be reached when $\xi_u$ is the uniform measure in $\mathcal{X}$ among the class $\mathcal{C}$ which contains all the measures $\xi_u$ satisfying that $\boldsymbol{H}^S(\xi_u)$ is not a negative matrix. In the sense of approximate designs, a UD should be selected as $d_u$ of the design $d$, especially when $\|h\|$ is large. The condition that $\boldsymbol{H}^S(\xi_u)$ is not a negative definite matrix in Theorem 2 can always be satisfied. See the following example.

**Example 2.** Let $\mathcal{X} = [-1, 1]$ and consider the location model $\boldsymbol{g}(x) = 1$. In this case, choose

$$\mathcal{K}(x, \omega) = \begin{cases} x + 2 - 3(x - x^2/2 + 7/2)(\omega - \omega^2/2 + 7/2)/20, & \text{if } x \le \omega, \\ \omega + 2 - 3(x - x^2/2 + 7/2)(\omega - \omega^2/2 + 7/2)/20, & \text{if } x > \omega, \end{cases}$$

as the reproducing kernel for $\mathcal{H}$. Moreover, choose $(-1, 1, 0, 0)^T$ as $d_f$ where $d_1 = (-1, 1)^T$ and two center points are chosen. Let the probability distribution function $\xi_u$ be $3\omega^2/2$ and $5\omega^4/2$, respectively. Then the corresponding values of $\boldsymbol{H}^S(\xi_u)$ are $n_u/50$ and $8n_u/105$, respectively. Both of them are larger than zero since $n_u > 0$. If $\xi_u$ is the uniform measure in $\mathcal{X}$, $\boldsymbol{H}^S(\xi_u)$ equals 0.

From Theorem 2, a composite design combining a two-level OA $d_1$, center points $d_0$ and a UD $d_2$, is more robust than any other composite design with the same $d_0$ and $d_1$ in the sense of approximate designs. Therefore, an OUCD combining a two-level OA $d_1$, a UD $d_2$ and center points $d_0$ has a similar property, i.e., OUCDs are robust.

### 3.2. Estimation efficiency

We now compare OUCDs with other composite designs with respect to estimation efficiency for the following second-order model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \varepsilon, \tag{7}$$

where $\beta_0$, $\beta_i$, $\beta_{ii}$ and $\beta_{ij}$ are the coefficients of intercept, linear, quadratic and bilinear terms, respectively, and $\epsilon \sim N(0, \sigma^2)$. For an $n$-run design, let $\boldsymbol{X}$ be the design matrix and $\boldsymbol{M} = \boldsymbol{X}^T\boldsymbol{X}/n$ be the information matrix. Let $\xi_{\text{opt}}$ be the continuous $D$-optimal design over $\mathcal{X} = [-1, 1]^k$ and denote $\text{Max}D_k = |\boldsymbol{M}(\xi_{\text{opt}})|$. The $D$-efficiency and $D_s$-value of the subset of the second-order model (7) for a $k$-factor, $n$-run design $d$ are respectively defined by

$$D_{\text{eff}}(d) = \left(\frac{|\boldsymbol{M}(d)|}{|\boldsymbol{M}(\xi_{\text{opt}})|}\right)^{1/p} = \frac{1}{n}\left(\frac{|\boldsymbol{X}^T\boldsymbol{X}|}{\text{Max}D_k}\right)^{1/p}, \tag{8}$$

$$D_s(d) = \frac{1}{n}\left(\frac{|\boldsymbol{X}^T\boldsymbol{X}|}{|\boldsymbol{X}_{(s)}^T\boldsymbol{X}_{(s)}|}\right)^{1/|s|} = \frac{1}{n}|\boldsymbol{X}_s^T\boldsymbol{X}_s - \boldsymbol{X}_s^T\boldsymbol{X}_{(s)}(\boldsymbol{X}_{(s)}^T\boldsymbol{X}_{(s)})^{-1}\boldsymbol{X}_{(s)}^T\boldsymbol{X}_s|^{1/|s|}, \tag{9}$$

where $p = (k+1)(k+2)/2$ is the number of parameters, $\boldsymbol{X}_s$ and $\boldsymbol{X}_{(s)}$ are the submatrices of $\boldsymbol{X}$ corresponding to the parameters in $s$ and not in $s$, respectively, and $|s|$ is the number of the parameters in $s$. Kiefer (1961) gave the approximate $D$-optimal and $D_s$-optimal designs over $\mathcal{X}$ for the second-order model (7). The optimal $D_s$-values with respect to linear $L$, bilinear $B$ and quadratic $Q$ terms are 1, 1 and 0.25, respectively. Then, the $D_s$-efficiency can be calculated by $D_{L,\text{eff}}(d) = D_L(d)$, $D_{B,\text{eff}}(d) = D_B(d)$ and $D_{Q,\text{eff}}(d) = 4D_Q(d)$.

Unfortunately, the $D$-efficiency and $D_s$-efficiency of the subset of the second-order model (7) do not have the explicit expressions in terms of OUCD. Then, we consider the $T$-criterion which represents the average of the diagonal elements of the information matrix to measure the estimation efficiency of OUCDs for the second-order model (7). Define $T$-efficiency for a $k$-factor, $n$-run design $d$ by

$$T_{\text{eff}}(d) = \frac{1}{p}\text{tr}(\boldsymbol{M}) = \frac{1}{np}\text{tr}(\boldsymbol{X}^T\boldsymbol{X}), \tag{10}$$

where $p = (k+1)(k+2)/2$. Silvey and Titterington (1974) showed that the $T$-optimal design $\tau$ only has the cube points over $\mathcal{X}$, and the maximum value of $T$-criterion is 1.

**Proposition 1.** *The $T$-efficiencies of a* CCD *and an* OACD *are*

$$T_{\text{eff}}(\text{CCD}) = \frac{2}{N_{k1}}\left[n_1 + 2k + n_0 + 2k(n_1+2) + \frac{k(k-1)}{2}n_1\right], \tag{11}$$

$$T_{\text{eff}}(\text{OACD}) = \frac{2}{N_{k2}}\left[n_1 + n_2 + n_0 + 2k\left(n_1 + \frac{2}{3}n_2\right) + \frac{k(k-1)}{2} \times \left(n_1 + \frac{4}{9}n_2\right)\right], \tag{12}$$

*respectively, where $N_{k1} = (k+1)(k+2)(n_1+2k+n_0)$ and $N_{k2} = (k+1)(k+2)(n_1+n_2+n_0)$. Moreover, let $d_2$ of the* OUCD *be a uniform* $\text{U}(n_2, 3^k)$*. The $T$-efficiency of the* OUCD *has the lower and upper bounds,*

$$\delta_{\text{eff}}(\text{OUCD}) = \frac{2}{N_{k2}}\left[n_1 + n_2 + n_0 + 2k\left(n_1 + \frac{2}{3}n_2\right) + \frac{k(k-1)}{2} \times \left(n_1 + \frac{1}{3}n_2\right)\right], \tag{13}$$

$$\gamma_{\text{eff}}(\text{OUCD}) = \frac{2}{N_{k2}}\left[n_1 + n_2 + n_0 + 2k\left(n_1 + \frac{2}{3}n_2\right) + \frac{k(k-1)}{2} \times \left(n_1 + \frac{2}{3}n_2\right)\right], \tag{14}$$

*respectively.*

The proof of Proposition 1 is straightforward and we omit it. It is obvious that $T_{\text{eff}}(\text{OACD})$ and $\delta_{\text{eff}}$ (OUCD) are larger than $T_{\text{eff}}$ (CCD) in Proposition 1 when all types of composite designs have the same $n_0$, $n_1$ and $n_2$ with $n_2 > 3$. Then, the OUCD has a better $T$-efficiency than the CCD. And $T_{\text{eff}}(\text{OACD})$ falls in $[\delta_{\text{eff}}(\text{OUCD}), \gamma_{\text{eff}}(\text{OUCD})$. From (11)–(14), all of $T_{\text{eff}}(\text{CCD})$, $T_{\text{eff}}(\text{OACD})$, $\delta_{\text{eff}}(\text{OUCD})$ and $\gamma_{\text{eff}}(\text{OUCD})$ decrease as $n_0$ increases. Given that $n_0 = 0$, $\delta_{\text{eff}}(\text{OUCD})$ and $\gamma_{\text{eff}}(\text{OUCD})$ monotonically increase in regard to the ratio $r = n_1/n_2$. Moreover, the difference between $T_{\text{eff}}(\text{OACD})$ and $\delta_{\text{eff}}(\text{OUCD})$ or $\lambda_{\text{eff}}$ (OUCD) decreases along with the increase of $r$ for each $k$. An OUCD with a large $r$ may be chosen in the sense of $T$-efficiency. Therefore, given $n_1$, we choose a small $n_2$ ensuring that the OUCD is a second-order design.

### 3.3. Space-filling property

In this subsection, the space-filling property of an OUCD under the maximin distance criterion is discussed. Define $L_1(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{i=1}^{k} |x_{1i} - x_{2i}|$ as the $L_1$-distance of any two rows $\boldsymbol{x}_1 = (x_{11}, \ldots, x_{1k})$ and $\boldsymbol{x}_2 = (x_{21}, \ldots, x_{2k})$ in a $k$-factor design $d$. Define the $L_1$-distance of $d$ as $L_1(d) = \min\{L_1(\boldsymbol{x}_1, \boldsymbol{x}_2) : \boldsymbol{x}_1 \neq \boldsymbol{x}_2, \boldsymbol{x}_1, \boldsymbol{x}_2 \in d\}$. The maximin OUCD is the optimal one under the maximin distance criterion. Without loss of generality, the center points are ignored in this subsection since $L_1(\text{OUCD}) = 0$ if it has more than one center point.

If a GLP set $D$ is chosen as the $d_2$ of an OUCD, there may exist a run whose elements are ones as in $d_1$. For avoiding the repeated runs in the OUCD, one method is to use the leave-one-out GLP set, another method is the linear level permutation method. For an $N \times k$ GLP set $D$, denote $D_{\boldsymbol{u}} = D + \mathbf{1}_N \boldsymbol{u} = \{\boldsymbol{x}_i + \boldsymbol{u}, \boldsymbol{x}_i \in D\} \pmod N$, where $\boldsymbol{u} = (u_1, \ldots, u_k)$, $u_j \in \{1, \ldots, N\}, j = 1, \ldots, k$, $\mathbf{1}_N$ is the $N \times 1$ column vector of ones and $D_{\boldsymbol{u}}$ is modified by replacing 0 by $N$. Zhou and Xu (2015) showed that any linear level permutation of a GLP set $D_{\boldsymbol{u}}$ does not decrease the $L_1$-distance. When $u_1 = \cdots = u_k = u$, $D_{\boldsymbol{u}}$ becomes the simple linear level permutation $D + u\boldsymbol{J}_{N\times k}$, we denote the corresponding transformed design $D^*_{\boldsymbol{u}}$ by $D^*_u$ and one can quickly find a design with a better $L_1$-distance. For an OUCD in $\mathcal{X}$, consider the simple linear level permutation for a GLP set $D$ such that $D^*_u$ is used for its $n_2$-run $d_2$ where $u = 1, \ldots, n_2$, and denote the permuted design by $\text{OUCD}_u$.

**Example 3.** Consider the OUCD shown in Table 1 after deleting the center points, and denote the 5-run GLP set in Example 1 by $D$. Since both $d_1$ and $d_2$ in the OUCD have a row vector of ones, $L_1(\text{OUCD}) = 0$. Consider the simple linear level permutation of $D$. When $u = 2, 3, 4$, $L_1(\text{OUCD}_u) = 0.5$.

**Proposition 2.** *If the $d_1$ in an OUCD is the two-level full factorial and $d_2$ is a GLP set, then $L_1(\text{OUCD}_u) = L_1(\text{OUCD}_\upsilon)$, where $\upsilon = n_2 + 1 - u$.*

From Proposition 2, $L_1(\text{OUCD}_u)$ is symmetric in regard to $u = (n_2+1)/2$ if $n_2$ is odd or $u = n_2/2$ if $n_2$ is even. Therefore, we only need to calculate $(n_2 + 1)/2$ or $n_2/2 L_1$-distances of the $\text{OUCD}_u$ and choose the best $\text{OUCD}_u$ under the maximin distance criterion. Moreover, different column alignments of $d_1$ and $d_2$ may lead to different $L_1$-distances. We directly have the following result and omit its proof.

**Proposition 3.** *If the $d_1$ in an OUCD is the full factorial or the $\text{OA}(2^{k-1}, 2^k, k-1)$ defined by $I = 1 \cdots k$, then the $L_1$-distance is invariant among all column alignments.*

It is obvious that $L_1(\text{OUCD})$ is not larger than either $L_1(d_1)$ or $L_1(d_2)$ of an OUCD. It is not easy to find a common tight upper bound of $L_1(\text{OUCD})$ for any OUCD, and some maximin two-level designs can be distinguished as $d_1$ and their $L_1$-distances can be the upper bounds of the resulting OUCDs.

**Lemma 1.** *If the $d_1$ in an OUCD is the two-level full factorial, $L_1(\text{OUCD})$ has the upper bound 2. And if $d_1$ is a regular $2^{k-1}$ or $2^{k-2}$ design, $L_1(\text{OUCD})$ has the upper bound 4 for every $k \geq 3$ or every $k \geq 5$, respectively.*

The proof of Lemma 1 is straightforward and omitted here. We now provide some conditions under which the OUCDs are the maximin distance designs as follows.

**Theorem 3.** *Let the $d_1$ in an OUCD be an $\text{OA}(2^k, 2^k, k)$ or $\text{OA}(2^{k-1}, 2^k, k-1)$, and $d_2$ be an $n_2 \times k$ leave-one-out GLP set with $k = \phi(n_2 + 1)$. Let $p_1$ and $p_2$ be two odd primes with $p_1 < p_2$ and $t$ be a positive integer. The OUCD is a maximin distance design if it satisfies any of the following conditions.*

(i) *Let $n_2 = p_1^t - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $3 \leq p_1 \leq 5$ and $t \geq 2$; or $p_1 \geq 7$. Otherwise, $p_1 = 3$ and $t \geq 3$; $5 \leq p_1 \leq 7$ and $t \geq 2$; or $p_1 \geq 11$.*
(ii) *Let $n_2 = 2p_1 - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $p_1 \geq 7$. Otherwise, $p_1 \geq 11$.*
(iii) *Let $n_2 = p_1 p_2 - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $p_1$ and $p_2$ are free. Otherwise, $p_1 p_2 > 15$.*
(iv) *Let $n_2 = 2^t - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $t \geq 4$. Otherwise, $t \geq 5$.*

Theorem 3 discusses the cases of $k = \phi(n_2 + 1)$. We now consider the cases of $k = \phi(n_2 + 1)/2$. Let $h_1, \ldots, h_{\phi(n_2+1)}$ be the integers coprime to $n_2 + 1$ with $1 = h_1 < \cdots < h_{\phi(n_2+1)} \leq n_2$.

**Theorem 4.** *Let $k = \phi(n_2 + 1)/2$, $p_1$ and $p_2$ be two odd primes with $p_1 < p_2$ and $t$ be a positive integer. Let the $d_1$ in an OUCD be an $\text{OA}(2^k, 2^k, k)$ or $\text{OA}(2^{k-1}, 2^k, k-1)$, and $d_2$ be an $n_2 \times k$ leave-one-out GLP set generated by $\boldsymbol{h} = (h_1, \ldots, h_k)$ or $(h_{k+1}, \ldots, h_{\phi(n_2+1)})$. The OUCD is a maximin distance design if it satisfies any of the following conditions.*

(i) *Let $n_2 = p_1^t - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $p_1 = 3$ and $t \geq 3$; $5 \leq p_1 \leq 7$ and $t \geq 2$; or $p_1 \geq 11$. Otherwise, $p_1 = 3$ and $t \geq 3$; $5 \leq p_1 < 19$ and $t \geq 2$; or $p_1 \geq 19$.*
(ii) *Let $n_2 = 2p_1 - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $p_1 \geq 11$. Otherwise, $p_1 \geq 19$.*
(iii) *Let $n_2 = p_1 p_2 - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $p_1 p_2 > 15$. Otherwise, $p_1 = 3$ and $p_2 \geq 11$; or $p_1 > 3$.*
(iv) *Let $n_2 = 2^t - 1$. If $d_1$ is the $\text{OA}(2^k, 2^k, k)$, $t \geq 5$. Otherwise, $t \geq 6$.*

The proof of Theorem 4 is similar to that of Theorem 3 and thus omitted. Moreover, some maximin OUCDs which have maximin regular $2^{k-2}$ designs as $d_1$ can also be constructed, and their $L_1$-distances reach the upper bound 4.

**Table 3**
The $D$-efficiencies of OACD, OUCD-I, OUCD-II, CCD, DSCD and UD.

| $k$ | $d_1$ | Generators | OACD | | OUCD-I | | OUCD-II | | CCD | DSCD | UD |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $d_2$ | $D_{\text{eff}}$ | $n_2$ | $D_{\text{eff}}$ | $n_2$ | $D_{\text{eff}}$ | $D_{\text{eff}}$ | $D_{\text{eff}}$ | $D_{\text{eff}}$ |
| 4 | $2^4$ | – | OA(9, $3^4$) | 0.931 | 9 | 0.891 | 5 | 0.818 | 0.936 | 0.891 | 0.324 |
| 5 | $2_V^{5-1}$ | E = ABCD | OA(18, $3^5$) | 0.953 | 18 | 0.944 | 7 | 0.779 | 0.869 | 0.846 | 0.257 |
| 6 | $2_{VI}^{6-1}$ | F = ABCDE | OA(18, $3^6$) | 0.966 | 18 | 0.948 | 7 | 0.777 | 0.868 | 0.837 | 0.241 |
| 7 | $2_{VII}^{7-1}$ | G = ABCDEF | OA(18, $3^7$) | 0.945 | 18 | 0.937 | 11 | 0.752 | 0.853 | 0.821 | 0.239 |
| 8 | $2_V^{8-2}$ | G = ABCDE, H = ABCF | OA(27, $3^8$) | 0.963 | 27 | 0.957 | 11 | 0.754 | 0.842 | 0.817 | 0.207 |
| 9 | $2_V^{9-2}$ | H = ABCDE, J = ABCFG | OA(27, $3^9$) | 0.950 | 27 | 0.946 | 11 | 0.727 | 0.829 | 0.802 | 0.207 |
| 10 | $2_V^{10-3}$ | H = ABCDE, J = ABCFG, K = ABDF | OA(27, $3^{10}$) | 0.952 | 27 | 0.948 | 11 | 0.738 | 0.830 | 0.808 | 0.190 |
| 11 | $2_V^{11-4}$ | H = ABCDE, J = ABCFG, K = ABDF, L = ACEG | OA (27, $3^{11}$) | 0.954 | 27 | 0.949 | 13 | 0.752 | 0.828 | 0.811 | 0.173 |
| 12 | OA(128, $2^{12}$, 4) | | OA(27, $3^{12}$) | 0.953 | 27 | 0.946 | 13 | 0.757 | 0.825 | 0.811 | 0.159 |

**Theorem 5.** *Consider any of the maximin* OUCDs *with* $d_1$ *being regular* $2^{k-1}$ *designs in* Theorems 3 *and* 4*. Let distinct* $\kappa_1, \ldots, \kappa_\ell \in \{1, \ldots, k-2\}$ *and* $2 \le \ell < k-2$*. Replacing* $d_1$ *by the regular* $2^{k-2}$ *design defined by* $(k-1) = 1 \cdots (k-2)$ *and* $k = \kappa_1 \cdots \kappa_\ell$*, the corresponding* OUCD *is also a maximin* OUCD.

Theorems 3–5 consider the $d_1$ as the two-level full factorials, regular $2^{k-1}$ and $2^{k-2}$ designs. However, they may still have too many runs to use in practice, and some regular $2^{k-m}$ designs with $m \ge 3$ may be considered for $d_1$. Small number of runs in $d_1$ often results in large $L_1$-distance of the corresponding OUCD when $d_2$ is fixed. Unfortunately, there is no clear relationship between the $L_1$-distance of the maximin regular $2^{k-m}$ design with $m \ge 3$ and its resolution. For searching a maximin OUCD with given $k$ and $m \ge 3$, we first find a maximin $2^{k-m}$ design among all defining words to form $d_1$. Let its $L_1$-distance be the upper bound of the $L_1$-distance of an OUCD whose $d_1$ is the regular $2^{k-m}$ design. Then choose a suitable leave-one-out GLP set as $d_2$ such that the corresponding OUCD reaches the upper bound. For instance, the 9-factor OUCD with $m = 3$ in Table 2 is a maximin OUCD. With the increasing of $m$, it is hard to find a UD $d_2$ such that $L_1$(OUCD) reaches the upper bound. However, when $d_1$ is the regular two-level saturated OA, Mukerjee and Wu (1995) summarized that the number of different elements between any two rows of the two-level saturated OA is $n_1/2$. Therefore, the $L_1$-distance of the saturated OA equals $n_1$, which is beneficial to be an upper bound for OUCDs.

## 4. Comparisons with other types of designs

### 4.1. D-Efficiency and T-efficiency

We now compare OUCDs with the OACDs, CCDs, DSCDs and UDs for $k = 4, \ldots, 12$ in terms of the $D$-efficiency and the $D_s$-efficiency, $s = L, B, Q$. Suppose $n_0 = 0$ for each type of those composite designs. For OUCDs, we keep the two-level OA $d_1$ as that in Table 1 of Zhou and Xu (2017), and have two choices of $d_2$, i.e., one is a three-level uniform U($n_2$, $3^k$), the other is an $n_2$-level GLP set, the two corresponding OUCDs are denoted by OUCD-I and OUCD-II, respectively. Let OUCD-I have the same number of runs as the corresponding OACD has, and OUCD-II select the smallest GLP set such that it is a second-order design. The UD is constructed under the uniformity criterion MD directly, and it has the same number of runs as the OACD has for each $k$. Both the $d_2$ of OUCD-I and UD are constructed by the R package named UniDOE by Zhang et al. (2018a) with arguments *init* = “*rand*", *crit* = “*MD*2" and *maxiter* = 100. All GLP sets are constructed by optimal generator vectors under MD. All of the UDs and the $d_2$ in each of the two types of OUCDs are shown in the supplementary material.

Table 3 illustrates the $D$-efficiency of every design for $k = 4, \ldots, 12$. It shows that each OUCD-I has a larger $D$-efficiency than the CCD and DSCD when $k > 4$, and both the OUCD-I and OUCD-II have larger $D$-efficiencies than the UD. Moreover, $D$-efficiencies of OUCD-I are close to that of OACDs. The reason is that each OUCD-I has many runs on the boundary of $\mathcal{X}$ which is like the optimal design $\xi_{\text{opt}}$. Next, we compare the OUCDs with the OACDs, DSCDs, CCDs and UDs in Table 3 in terms of $D_s$-efficiency, $s = L, B, Q$. It is shown that the $D_L$-efficiencies of OUCD-I are larger than that of the CCDs and UDs for $k = 4, \ldots, 12$, and than the OACDs when $k = 4, \ldots, 9$. Moreover, each OUCD-II also has a larger $D_L$-efficiency than the CCD and UD. For $D_Q$-efficiency, each OUCD-I has a larger one than the CCD except when $k = 4, 5$, and the DSCD except when $k = 4$. For $D_B$-efficiency, the OUCD-II has a larger one than other types of designs for each $k$, and each OUCD-I has a larger one than the DSCD except when $k = 8$ and the UD. In conclusion, OUCDs perform well in terms of $D$-efficiency, $D_L$-efficiency and $D_B$-efficiency.

In addition, we also compare the OUCD with the OACD, CCD, DSCD and UD in terms of $T$-efficiency for every $k = 4, \ldots, 12$. Both $\delta_{\text{eff}}$(OUCD-I) and $T_{\text{eff}}$(OUCD-II) are larger than $T_{\text{eff}}$(CCD) and $T_{\text{eff}}$(UD) for each $k$. Each $T_{\text{eff}}$(OUCD-II)

**Table 4**
The $D$-efficiencies among all possible projection designs of the OUCD-I and OUCD-II.

| $k$ | 9-factor OUCDs | | | | | | 10-factor OUCDs | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | OUCD-I | | | OUCD-II | | | OUCD-I | | | OUCD-II | | |
| | mean | min | max | mean | min | max | mean | min | max | mean | min | max |
| 4 | 0.983 | 0.981 | 0.985 | 0.778 | 0.777 | 0.779 | 0.952 | 0.948 | 0.955 | 0.756 | 0.581 | 0.784 |
| 5 | 0.991 | 0.989 | 0.993 | 0.775 | 0.774 | 0.775 | 0.969 | 0.965 | 0.972 | 0.756 | 0.639 | 0.790 |
| 6 | 0.992 | 0.990 | 0.993 | 0.772 | 0.772 | 0.773 | 0.977 | 0.974 | 0.980 | 0.756 | 0.679 | 0.795 |
| 7 | 0.989 | 0.987 | 0.990 | 0.769 | 0.769 | 0.770 | 0.980 | 0.978 | 0.982 | 0.756 | 0.707 | 0.798 |
| 8 | 0.983 | 0.981 | 0.983 | 0.767 | 0.767 | 0.767 | 0.979 | 0.978 | 0.981 | 0.755 | 0.727 | 0.801 |
| 9 | – | – | – | – | – | – | 0.976 | 0.976 | 0.977 | 0.754 | 0.742 | 0.802 |

**Table 5**
The $\rho^2$ and $\varrho$ of the OUCDs in Table 3.

| $k$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|
| $\rho^2$(OUCD-I) | 0 | 0 | 0 | 0 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $\rho^2$(OUCD-II) | 0.003 | 0.004 | 0.001 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| $\varrho$(OUCD-II) | 0.012 | 0.021 | 0.006 | 0.004 | 0.004 | 0.001 | 0.001 | 0.001 | 0.001 |

is larger than $T_{\rm eff}$(OACD) except when $k = 7$, and the $T_{\rm eff}$(OUCD-I) and $T_{\rm eff}$(OACD) for each $k$ are close to each other. Therefore, OUCDs also have good performances in terms of $T$-efficiency.

It should be mentioned that each OUCD-II listed in Table 3 is not a maximin OUCD. One can also use a leave-one-out GLP as $d_2$ which has a smaller number of runs than that of OUCD-I to construct a maximin OUCD when $d_1$ is a $2^{k-m}$ design with maximin distance. For example, when $d_1$ is the $2^{9-2}$ in Table 3, $d_1$ is a maximin distance design and its $L_1$-distance equals 4. Let $d_2$ be a 18-run 9-factor leave-one-out GLP set with the generator vector $h = (1, \dots, 9)$. The composite design $d$ combining the $d_1$ and $d_2$ is a maximin OUCD with $L_1$-distance 4, and its $D$-efficiency $D_{\rm eff}(d) = 0.832$, $T$-efficiency $T_{\rm eff}(d) = 0.902$ and $D_s$-efficiencies $D_L(d) = 0.922$, $D_B(d) = 0.877$ and $D_Q(d) = 0.079$. Then $d$ has a larger $D$-efficiency than the OUCD-II, CCD, DSCD and UD in Table 3, and has a larger $D_Q$-efficiency than the OUCD-II and DSCD and also has a larger $D_B$-efficiency than the OUCD-I and DSCD. Therefore, the maximin OUCD may have appealing space-filling property and high estimation efficiency.

Furthermore, consider the projection property of OUCDs. For a fair comparison, let the OUCD-I and OUCD-II have the same $d_1$ part and the run size $n_2$ as the corresponding OACD in Table 5 of Zhou and Xu (2017) has for each $k = 9, 10$. Here, the 9-factor $d_2$ is constructed by the GLP method with power generator (Fang et al., 2018) under MD criterion. The $d_2$ in each of the two types of OUCDs is shown in the supplementary material. We calculate the mean, maximum and minimum $D$-efficiencies among all possible projection designs of each OUCD. The results are shown in Table 4. It can be easily seen that all the projection designs of the OUCD-I have appealing $D$-efficiencies. For the mean, maximum and minimum $D$-efficiencies, the OUCD-I performs similarly as the OACD in Table 5 of Zhou and Xu (2017) for each $k = 9, 10$. Then, OUCD-I is a better choice among OUCDs in terms of projection property in $D$-efficiency.

### *4.2. Orthogonality*

The orthogonality of OUCD is discussed in this subsection. Owen (1994) proposed a criterion $\rho^2$ to measure the orthogonality of a $k$-factor design with respect to pairwise correlation between columns of the design, i.e.,

$$\rho^2 = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \rho_{ij}^2, \tag{15}$$

where $\rho_{ij}$ is the linear correlation coefficient between the $i$th and $j$th columns of the design with $i \neq j$. The $\rho^2$ of each OUCD in Table 3 is shown in Table 5. It can be seen that all the values of $\rho^2$ are equal to 0 or close to 0, which means that the OUCDs are nearly orthogonal and have small linear correlations between columns. For example, $\rho^2$(OUCD-I) equals 0 for each $k = 4, \dots, 7$ since the $d_2$ of each OUCD-I is not only uniform but also orthogonal. The following result gives an upper bound of $\rho^2$.

**Proposition 4.** *If $n_2$ is an odd prime, let the $d_2$ in an* OUCD *be an $n_2$-run* GLP *set, then the upper bound of $\rho^2$(OUCD) is $\varrho(\text{OUCD}) = \alpha^2/\eta^2$, where $\alpha = 2-n_2+4n_2(n_2-2)/[3(n_2-1)]$ and $\eta = (n_2^2+n_2)/(3n_2-3)+n_1$. Let $d_2$ be an $n_2$-run leave-one-out* GLP *set if $n_2+1$ is an odd prime, then $\rho^2(\text{OUCD}) \le \varrho(\text{OUCD}) = \theta^2/\vartheta^2$, where $\theta = n_2[(n_2+1)(n_2-1)-6]/[3(n_2-1)^2]$ and $\vartheta = 2n_2(2n_2-1)/[3(n_2-1)] + n_1 - n_2$.*

The $\varrho$(OUCD) of each OUCD-II listed in Table 3 is also shown in Table 5. It can be seen that those values of upper bound are also very small. In addition, the linear level permutation is a useful technique for GLP sets to improve their $\rho^2$ as well as that of the OUCDs with the GLP sets as $d_2$. If $\boldsymbol{u} = (u, \dots, u)$, we have the following result.

**Proposition 5.** *Let the $d_1$ in an OUCD be a two-level OA and $d_2$ be an $n_2$-run GLP set. Then, $\rho^2(\mathrm{OUCD}_u) = \rho^2(\mathrm{OUCD}_\upsilon)$, where $u+\upsilon = n_2+1$ and $u \in \{1, \dots, n_2\}$. If $n_2$ is even, $\rho^2(\mathrm{OUCD}_\varsigma) = \rho^2(\mathrm{OUCD}_\nu)$, where $\varsigma + \nu = n_2/2+1$ and $\varsigma \in \{1, \dots, n_2/2\}$.*

**Example 4.** Consider the OUCD shown in Example 3, and replace the $d_2$ by the 10-run GLP set with $\boldsymbol{h} = (1, 3, 7)$, it holds that $\rho^2(\mathrm{OUCD}_1) = \rho^2(\mathrm{OUCD}_5) = \rho^2(\mathrm{OUCD}_6) = \rho^2(\mathrm{OUCD}_{10}) = 0.014$, $\rho^2(\mathrm{OUCD}_2) = \rho^2(\mathrm{OUCD}_4) = \rho^2(\mathrm{OUCD}_7) = \rho^2(\mathrm{OUCD}_9) = 0.003$ and $\rho^2(\mathrm{OUCD}_3) = \rho^2(\mathrm{OUCD}_8) = 0.026$. Then $\rho^2(\mathrm{OUCD}_2) < \rho^2(\mathrm{OUCD}_1)$ which means that the simple linear level permutation may decrease $\rho^2$ of OUCDs.

## 5. Conclusions

This paper proposes a new type of composite designs, orthogonal uniform composite designs (OUCDs), which combine orthogonal arrays and uniform designs. OUCDs have more flexible numbers of runs than OACDs, CCDs and DSCDs. It is shown that OUCDs are more robust than other types of composite designs. The space-filling property of OUCDs under the maximin distance criterion is discussed, and construction methods for the $k$-factor maximin OUCD whose $d_1$ part is the two-level full factorial, regular $2^{k-1}$ or $2^{k-2}$ design are provided. Moreover, for the estimation efficiency for the second-order model (7), OUCDs have larger $D$-efficiencies than CCDs, DSCDs and UDs, larger $T$-efficiencies than CCDs and UDs, and have larger $D_B$-efficiencies than other types of composite designs. Moreover, OUCDs are nearly orthogonal. They can also be used to perform multiple analysis for cross, i.e., the $d_1$ part and $d_2$ part can be analyzed separately.


## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 11871288 and 11771220), National Ten Thousand Talents Program, Tianjin Development Program for Innovation and Entrepreneurship, Tianjin"131" Talents Program and Ph.D. Candidate Research Innovation Fund of Nankai University. The authors would like to thank Prof. Dennis K. J. Lin, Ms. Chun-Yan Wang, the editor and the two anonymous referees for their valuable comments and suggestions.


## Appendix A. Proofs

The following lemma from linear algebra is used to prove Theorems 1 and 2, and one can refer to Wang et al. (2006) for the detailed proof.

**Lemma A.1** (*Wang et al., 2006*)**.** *Let $\lambda_1(\boldsymbol{A}), \dots, \lambda_N(\boldsymbol{A})$ be $N$ eigenvalues of an $N \times N$ square matrix $\boldsymbol{A}$ which are listed in descending order, particularly where $\lambda_1(\boldsymbol{A})$ is the maximum eigenvalue and $\lambda_N(\boldsymbol{A})$ is the minimum eigenvalue. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two $N \times N$ Hermite matrices. We have the following results.*

(a) *If $\boldsymbol{A} \le \boldsymbol{B}$, then $\lambda_i(\boldsymbol{A}) \le \lambda_i(\boldsymbol{B})$.*
(b) *$\lambda_i(\boldsymbol{A}) + \lambda_N(\boldsymbol{B}) \le \lambda_i(\boldsymbol{A}+\boldsymbol{B}) \le \lambda_i(\boldsymbol{A}) + \lambda_1(\boldsymbol{B})$.*
(c) *Additionally, $\boldsymbol{B}$ is positive semidefinite, then $\lambda_i(\boldsymbol{A}^2)\lambda_N(\boldsymbol{B}) \le \lambda_i(\boldsymbol{ABA}) \le \lambda_i(\boldsymbol{A}^2)\lambda_1(\boldsymbol{B})$.*
(d) *Additionally, $\boldsymbol{A}$ and $\boldsymbol{B}$ are positive semidefinite, then $\lambda_N(\boldsymbol{A})\lambda_i(\boldsymbol{B}) \le \lambda_i(\boldsymbol{AB}) \le \lambda_1(\boldsymbol{A})\lambda_i(\boldsymbol{B})$ and $\lambda_i(\boldsymbol{A})\lambda_N(\boldsymbol{B}) \le \lambda_i(\boldsymbol{AB}) \le \lambda_i(\boldsymbol{A})\lambda_1(\boldsymbol{B})$.*
(e) *Additionally, $\boldsymbol{A}$ and $\boldsymbol{B}$ are positive semidefinite, then $0 \le tr(\boldsymbol{AB}) \le \lambda_1(\boldsymbol{B})tr(\boldsymbol{A}) \le tr(\boldsymbol{A})tr(\boldsymbol{B})$.*

**Proof of Theorem 1.** According to Theorem 1 of Yue and Hickernell (1999), $L(d, h)$ in (3) has the upper bound, namely,

$$L(d,h) \le \sigma^2 \mathrm{tr}\left(\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{G}\right) + \lambda_{\max}(\boldsymbol{T})\,\|h\|^2. \tag{A.1}$$

Moreover, the upper bound (A.1) can be reached when $h = c\boldsymbol{v}_{\max}^T(\boldsymbol{T})\boldsymbol{G}^{1/2}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{k}$, $c \in \mathbb{R}$. Since $\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}$ and $\boldsymbol{G}$ are two $p \times p$ positive definite matrices and $\boldsymbol{X}^T\boldsymbol{X} \ge \boldsymbol{X}_f^T\boldsymbol{X}_f$, from Lemma A.1(a) and (e), it holds

$$\begin{aligned}\sigma^2\mathrm{tr}\left(\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{G}\right) &= \sigma^2\mathrm{tr}\left(\boldsymbol{G}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\right)\\ &\le \sigma^2\lambda_{\min}^{-1}\left(\boldsymbol{X}^T\boldsymbol{X}\right)\mathrm{tr}(\boldsymbol{G}) \le \sigma^2\lambda_{\min}^{-1}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)\mathrm{tr}(\boldsymbol{G}).\end{aligned} \tag{A.2}$$

The equality holds when the unfixed part $d_u$ is empty in $d$ and $\boldsymbol{X}_f^T\boldsymbol{X}_f = a\boldsymbol{I}_p$, $a \in \mathbb{R}$ and $a > 0$. In addition, from Lemma A.1(a)–(d), we have

$$\begin{aligned}\lambda_{\max}(\boldsymbol{T}) &= \lambda_{\max}\left(\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{G}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{K}\boldsymbol{X}\right)\\ &\le \lambda_{\max}\left(\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{G}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\right)\lambda_{\max}\left(\boldsymbol{X}^T\boldsymbol{K}\boldsymbol{X}\right)\\ &\le \lambda_{\max}(\boldsymbol{G})\lambda_{\min}^{-2}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)\left(\lambda_{\max}\left(\boldsymbol{X}_f^T\boldsymbol{K}_{ff}\boldsymbol{X}_f\right) + \lambda_{\max}(\boldsymbol{Q}) + \lambda_{\max}\left(\boldsymbol{X}_u^T\boldsymbol{K}_{uu}\boldsymbol{X}_u\right)\right)\end{aligned} \tag{A.3}$$

According to (A.1)–(A.3), we obtain (4) and the equality holds when the unfixed part $d_u$ is empty in $d$, $\boldsymbol{X}_f^T\boldsymbol{X}_f = a\boldsymbol{I}_p$, $\boldsymbol{G} = b\boldsymbol{I}_p$, $h = c\boldsymbol{v}_{\max}^T(\boldsymbol{T})\boldsymbol{G}^{1/2}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{k}$, where $a, b, c \in \mathbb{R}$, $a, b > 0$, and $\boldsymbol{v}_{\max}(\boldsymbol{T})$ is the (scaled) eigenvector of $\boldsymbol{T}$ corresponding to the maximum eigenvalue.

**Proof of Theorem 2.** Extending discrete distribution of $d_u$ to the continuous design measure $\xi_u$, we can obtain $\gamma(d_{\xi_u}, h)$ in (5) based on (4), the definitions of $\boldsymbol{H}_{fu}(\xi_u)$, $\boldsymbol{H}_{uf}(\xi_u)$, $\boldsymbol{H}_{uu}(\xi_u)$ and $\boldsymbol{H}^S(\xi_u)$. If $\boldsymbol{H}^S(\xi_u)$ is not a negative definite matrix, from Lemma A.1(a) and $\boldsymbol{H}_{uu}(\xi_u) \geq 0$ for any $\xi_u$, it holds that $\gamma(d_{\xi_u}, h) \geq \gamma_{\min}(\xi_u, h) = \sigma^2\lambda_{\min}^{-1}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)tr(\boldsymbol{G}) + \lambda_{\max}(\boldsymbol{G})\lambda_{\min}^{-2}\left(\boldsymbol{X}_f^T\boldsymbol{X}_f\right)\lambda_{\max}\left(\boldsymbol{X}_f^T\boldsymbol{K}_{ff}\boldsymbol{X}_f\right)\|h\|^2$, the equality holds when $\lambda_{\max}\left(\boldsymbol{H}^S(\xi_u)\right) = 0$ and $\lambda_{\max}(\boldsymbol{H}_{uu}(\xi_u)) = 0$. Moreover, $\lambda_{\max}\left(\boldsymbol{H}^S(\xi_u)\right) \geq 0$ when $\boldsymbol{H}^S(\xi_u)$ is not a negative definite matrix. If $\xi_u$ is the uniform measure in $\mathcal{X}$, $\boldsymbol{H}^S(\xi_u) = 0$ and $\boldsymbol{H}_{uu}(\xi_u) = 0$ since $\int_{\mathcal{X}} g_j(\boldsymbol{x})\mathcal{K}(\cdot, \boldsymbol{x})\mathrm{d}\boldsymbol{x} = 0, j = 1, \ldots, p$. Then, the equality holds when $\xi_u$ is the uniform measure in $\mathcal{X}$, and (6) follows.

**Proof of Proposition 2.** For convenience, we transform the design region $\mathcal{X}$ of an OUCD into $[1, n_2]^k$, i.e., change the levels $-1$ and 1 in $d_1$ to 1 and $n_2$ respectively, and the levels in $d_2$ are over $[1, n_2]$. In terms of the two-level full factorial design $d_1$ in the OUCD, the linear level permutation for $d_2$ does not change $L_1(d_1)$. Since $\boldsymbol{x}_{n_2} = n_2\boldsymbol{1}_k^T$ and $\boldsymbol{x}_i + \boldsymbol{x}_{n_2-i} = n_2\boldsymbol{1}_k^T$, $i = 1, \ldots, n_2 - 1$, for any $i$th row vector $\boldsymbol{x}_i \in d_2$, $i = 1, \ldots, n_2$, it holds that $\boldsymbol{x}_i + u\boldsymbol{1}_k^T \pmod{n_2} + (\boldsymbol{x}_{n_2-i} + \upsilon\boldsymbol{1}_k^T \pmod{n_2}) = (n_2 + 1)\boldsymbol{1}_k^T$, $u = 1, \ldots, n_2$. Then, for each $\boldsymbol{x}_j \in d_2$ with $j \neq i$, it holds that the $L_1$-distance between $\boldsymbol{x}_i + u\boldsymbol{1}_k^T \pmod{n_2}$ and $\boldsymbol{x}_j + u\boldsymbol{1}_k^T \pmod{n_2}$ is the same as the $L_1$-distance between $\boldsymbol{x}_{n_2-i} + \upsilon\boldsymbol{1}_k^T \pmod{n_2}$ and $\boldsymbol{x}_{n_2-j} + \upsilon\boldsymbol{1}_k^T \pmod{n_2}$. Therefore, $L_1(d_2 + u\boldsymbol{J}_{n_2\times k}) \geq L_1(d_2 + \upsilon\boldsymbol{J}_{n_2\times k})$. In addition, we have $L_1(d_2 + u\boldsymbol{J}_{n_2\times k}) \leq L_1(d_2 + \upsilon\boldsymbol{J}_{n_2\times k})$ when exchange $u\boldsymbol{1}_k^T$ and $\upsilon\boldsymbol{1}_k^T$. Then, $L_1(d_2 + u\boldsymbol{J}_{n_2\times k}) = L_1(d_2 + \upsilon\boldsymbol{J}_{n_2\times k})$.

Moreover, let $L_1(d_1, d_2)$ be the smallest $L_1$-distance between $\boldsymbol{x}_i \in d_2$ and $\boldsymbol{y}_\ell \in d_1$. If $L_1(d_1, d_2 + u\boldsymbol{J}_{n_2\times k}) = L_1(\boldsymbol{y}_\ell, \boldsymbol{x}_i + u\boldsymbol{1}_k^T) = z$, it holds that $L_1((n_2 + 1)\boldsymbol{1}_k^T - \boldsymbol{y}_\ell, \boldsymbol{x}_{n_2-i} + \upsilon\boldsymbol{1}_k^T) = z$. Then, $L_1(d_1, d_2 + u\boldsymbol{J}_{n_2\times k}) \geq L_1(d_1, d_2 + \upsilon\boldsymbol{J}_{n_2\times k})$ since $(n_2+1)\boldsymbol{1}_k^T - \boldsymbol{y}_\ell \in d_1$. Similarly, it is easy to know that $L_1(d_1, d_2 + u\boldsymbol{J}_{n_2\times k}) \leq L_1(d_1, d_2 + \upsilon\boldsymbol{J}_{n_2\times k})$. Therefore, $L_1(d_1, d_2 + u\boldsymbol{J}_{n_2\times k}) = L_1(d_1, d_2 + \upsilon\boldsymbol{J}_{n_2\times k})$, such that $L_1(\mathrm{OUCD}_u) = L_1(\mathrm{OUCD}_\upsilon)$.

**Proof of Theorem 3.** Let $k = \phi(n_2+1)$. And denote the greatest common divisor of $N$ and $h$ by $\gcd(N, h)$. Consider the case (i). If a $k$-factor OUCD has the OA$(2^k, 2^k, k)$ as $d_1$, then $L_1(d_1) = 2$. According to Theorem 4 of Zhou and Xu (2015), $L_1(d_2) = [(n_2+1)^2+p_1](p_1-1)/[2(n_2-1)p_1]$. If $\gcd(i, n_2+1) = 1$, $i = 1, \ldots, n_2$, $L_1(d_1, \boldsymbol{x}_i) = [(n_2+1)^2+p_1-4(n_2+1)](p_1-1)/[2(n_2-1)p_1]$ for each $\boldsymbol{x}_i \in d_2$. If $\gcd(i, n_2) = p_1^m$, $i = 1, \ldots, n_2$, $L_1(d_1, \boldsymbol{x}_i) = [(n_2+1)^2 + p_1^{2m+1} - 4(n_2+1)](p_1-1)/[2(n_2-1)p_1]$ for each $\boldsymbol{x}_i \in d_2$ and $m = 1, \ldots, t-1$. Then, $L_1(d_1, d_2) = [(n_2+1)^2 + p_1 - 4(n_2+1)](p_1-1)/[2(n_2-1)p_1]$ and $L_1(d_1, d_2) < L_1(d_2)$. It is obvious that $L_1(d_1, d_2)$ does not decrease as $t$ increases when we fix $p$. When $p_1 = 3, 5$ and $t \geq 2$; or $p_1 \geq 7$, $L_1(d_1, d_2) \geq 2$ and $L_1(\mathrm{OUCD}) = 2$ where the OUCD has the maximin $L_1$-distance of 2.

If a $k$-factor OUCD has the maximin OA$(2^{k-1}, 2^k, k-1)$ as $d_1$, $L_1(d_1) = 4$. Otherwise, $L_1(d_1) = 2$. If $d_2$ in the OUCD is an $n_2 \times k$ leave-one-out GLP set, then

$$L_1(d_1, d_2) = \begin{cases} \frac{[(n_2+1)^2+p_1-4(n_2+1)](p_1-1)+4p_1}{2(n_2-1)p_1}, & \text{if } (p_1-1)/2 \text{ is odd,} \\ \frac{[(n_2+1)^2+p_1-4(n_2+1)](p_1-1)}{2(n_2-1)p_1}, & \text{if } (p_1-1)/2 \text{ is even.} \end{cases}$$

Then, $L_1(d_2) > L_1(d_1, d_2)$. When $p_1 = 3$ and $t \geq 3$; $p_1 = 5, 7$ and $t \geq 2$; or $p_1 \geq 11$, $L_1(d_1, d_2) \geq 4$ such that the OUCD is a maximin OUCD since $L_1(\mathrm{OUCD})$ reaches the upper bound 4. The remaining cases can be similarly proved and their proofs are omitted.

**Proof of Theorem 5.** Consider the maximin OUCD combining the regular $2^{k-1}$ design and $n_2$-run $d_2$ under the case (i) in Theorem 3. Let us replace the regular $2^{k-1}$ design by the regular $2^{k-2}$ design defined by $(k-1) = 1\cdots(k-2)$ and $k = \kappa_1\cdots\kappa_\ell$ as $d_1$ and $d_1^{\mathrm{new}}$ be the OA$(2^k, 2^k, k)$, then $L_1(d_1, d_2) \geq L_1(d_1^{\mathrm{new}}, d_2)$. And $L_1(d_1^{\mathrm{new}}, d_2) = [(n_2+1)^2 + p_1 - 4(n_2+1)](p_1-1)/[2(n_2-1)p_1] \geq 4$ and $L_1(d_2) \geq 4$ such that the OUCD combining $d_1$ and $d_2$ is also a maximin OUCD. The remaining cases can be similarly proved and their proofs are omitted.

**Proof of Proposition 4.** If $d_2$ is an $n_2$-run GLP set with odd prime $n_2$, there does not exist the same level in each row except the last row. Let $\boldsymbol{x}^j$ be the $j$th column of the OUCD and $x_{ij}$ be its $i$th element. The mean of any column in the OUCD denoted by $\bar{x}_1$ is 0. Besides, any two columns of both $d_1$ and $d_0$ are linearly independent. Each of levels 1 and $-1$ appears $n_1/2$ times in $d_1$ and level 0 appears $n_0$ times in $d_0$. Due to the fact that each row of $d_2$ has distinct levels except the last row, then $\left[\sum_{i=1}^{n}(x_{i\ell} - \bar{x}_1)(x_{ij} - \bar{x}_1)\right]^2 = \left[\sum_{i=1}^{n_2}(x_{(i+n_0+n_1)\ell} - \bar{x}_1)(x_{(i+n_0+n_1)j} - \bar{x}_1)\right]^2 \leq \left\{1 + 2\sum_{i=0}^{(n_2-3)/2}[-1 + 4i/(n_2-1)][-1 + 2(2i+1)/(n_2-1)]\right\}^2 = \{2 - n_2 + 4n_2(n_2-2)/[3(n_2-1)]\}^2$, where $\boldsymbol{x}^\ell$ and $\boldsymbol{x}^j$ are any two distinct columns of the OUCD, and $n = n_1 + n_2 + n_0$. Also, $\sum_{i=1}^{n}(x_{i\ell} - \bar{x}_1)^2 = (\boldsymbol{x}^\ell)^T\boldsymbol{x}^\ell = (n_2^2 + n_2)/(3n_2 - 3) + n_1$ for $\ell = 1, \ldots, k$. Thus, for the linear correlation $\rho_{\ell j}$ between the distinct $\ell$th and $j$th columns, it holds that $\rho_{\ell j}^2 = \left[\sum_{i=1}^{n}(x_{i\ell} - \bar{x}_1)(x_{ij} - \bar{x}_1)/\sum_{i=1}^{n}(x_{i\ell} - \bar{x}_1)^2\right]^2 \leq \alpha^2/\eta^2$, where $\alpha = 2 - n_2 + 4n_2(n_2-2)/[3(n_2-1)]$ and $\eta = (n_2^2 + n_2)/(3n_2 - 3) + n_1$. Then, from (15), the upper bound $\varrho(\mathrm{OUCD})$ follows directly. If $d_2$ is an $n_2$-run leave-one-out

GLP set, there does not exist the row vector of ones. Similarly, we can prove the upper bound $\varrho(\mathrm{OUCD})$ and its proof is omitted.

**Proof of Proposition 5.** The $d_1$ in the OUCD is an OA so that we only focus on $d_2$, the GLP set $D$. Denote $\boldsymbol{x}_0 = \boldsymbol{x}_{n_2}$ where $\boldsymbol{x}_{n_2}$ is the last row in $D$. For each $i$th row vector $\boldsymbol{x}_i \in D$ with $i = 1, \ldots, n_2$, it holds that $\boldsymbol{x}_i + u\mathbf{1}_k^T \pmod{n_2} + (\boldsymbol{x}_{n_2-i} + \upsilon\mathbf{1}_k^T \pmod{n_2}) = (n_2+1)\mathbf{1}_k^T$ where the multiplication operation modulo $n_2$ is modified so that the result falls into $[1, n_2]$ by replacing 0 by $n_2$. Through the linear mapping $f$ for $D$, it holds that $\left(\boldsymbol{\chi}_i^{(u)}\right)^* + \left(\boldsymbol{\chi}_{n_2-i}^{(\upsilon)}\right)^* = 0$ where $\boldsymbol{\chi}_i^{(u)} = \boldsymbol{x}_i + u\mathbf{1}_{n_2}^T \pmod{n_2}$, $\upsilon = n_2 + 1 - u$ and $u \in \{1, \ldots, n_2\}$. Then, $D_u^*$ can be obtained by $-D_\upsilon^*$ under the row permutation. Since the permutation does not affect $\rho^2$, it holds that $\rho^2(\mathrm{OUCD}_u) = \rho^2(\mathrm{OUCD}_\upsilon)$.

If $n_2$ is even, there exists an $\ell$ such that $i + \ell \pmod{n_2} = n_2/2$ for any $i$, where both $i$ and $\ell$ belong to $\{1, \ldots, n_2\}$. Then, $\boldsymbol{x}_i + \varsigma\mathbf{1}_k^T \pmod{n_2} + (\boldsymbol{x}_\ell + \nu\mathbf{1}_k^T \pmod{n_2}) = (n_2+1)\mathbf{1}_k^T$. Similarly, it holds that $\rho^2(\mathrm{OUCD}_\varsigma) = \rho^2(\mathrm{OUCD}_\nu)$.

## Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jspi.2019.08.007.

## References


Box, G.E.P., Wilson, K.B., 1951. On the experimental attainment of optimum conditions. J. R. Stat. Soc. Ser. B Stat. Methodol. 13, 1–45.
Cheng, C.S., Deng, L.Y., Tang, B., 2002. Generalized minimum aberration and design efficiency for nonregular fractional factorial designs. Statist. Sinica 12, 991–1000.
Draper, N.R., Lin, D.K.J., 1990. Small response-surface designs. Technometrics 32, 187–194.
Fang, K.T., Liu, M.Q., Qin, H., Zhou, Y.D., 2018. Theory and Application of Uniform Experimental Designs. Springer, Singapore.
Gilmour, S., 2006. Response surface designs for experiments in bioprocessing. Biometrics 62, 323–331.
Hickernell, F.J., 1998. A generalized discrepancy and quadrature error bound. Math. Comput. 67, 299–322.
Johnson, M.E., Moore, L.M., Ylvisaker, D., 1990. Minimax and maximin distance design. J. Statist. Plann. Inference 26, 131–148.
Kiefer, J., 1959. Optimaml experimental designs (with discussion). J. R. Stat. Soc. Ser. B Stat. Methodol. 21, 272–319.
Kiefer, J., 1961. Optimal designs in regression problems II. Ann. Math. Statist. 32, 298–325.
Morris, M.D., 2000. A class of three-level experimental designs for response surface modeling. Technometrics 42, 111–121.
Mukerjee, R., Wu, C.F.J., 1995. On the existence of saturated and nearly saturated asymmetrical orthogonal arrays. Ann. Statist. 23, 2102–2115.
Owen, A., 1994. Controlling correlations in Latin hypercube samples. J. Amer. Statist. Assoc. 89, 1517–1522.
Silvey, S.D., Titterington, D.M., 1974. A Lagrangian approach to optimal design. Biometrika 61, 299–302.
Wang, S.G., Wu, M.X., Jia, Z.Z., 2006. Matrix Inequalities, (second ed. in Chinese) Science Press, Beijing.
Xu, H., Cheng, S.W., Wu, C.F.J., 2004. Optimal projective three-level designs for factor screening and interaction detection. Technometrics 46, 280–292.
Xu, H., Jaynes, J., Ding, X., 2014. Combining two-level and three-level orthogonal arrays for factor screening and response surface exploration. Statist. Sinica 24, 269–289.
Xu, H., Wu, C.F.J., 2001. Generalized minimum aberration for asymmetrical fractional factorial designs. Ann. Statist. 29, 1066–1077.
Yue, R.X., Hickernell, F.J., 1999. Robust designs for fitting linear models with misspecification. Statist. Sinica 9, 1053–1069.
Zhang, A.J., Li, H.Y., Quan, S.J., Yang, Z.B., 2018. UniDOE: Uniform design of experiments. R package version 1.0.2.
Zhang, X.R., Qi, Z.F., Zhou, Y.D., Yang, F., 2018b. Orthogonal-array composite design for the third-order models. Comm. Statist. Theory Methods 47, 3488–3507.
Zhou, Y.D., Fang, K.T., Ning, J.H., 2013. Mixture discrepancy for quasi-random point sets. J. Complexity 29, 283–301.
Zhou, Y.D., Xu, H., 2015. Space-filling properties of good lattice point sets. Biometrika 102, 959–966.
Zhou, Y.D., Xu, H., 2017. Composite designs based on orthogonal arrays and definitive screening designs. J. Amer. Statist. Assoc. 112, 1675–1683.